

งานวิจัยที่เกี่ยวข้องกับการรู้จำตัวอักษรภาษาไทยที่ผ่านมานุ่งเน้นในการเพิ่มความสามารถในระบบการรู้จำตัวอักษรภาษาไทยด้วยการพิจารณาโครงสร้างลักษณะเด่นของตัวอักษรที่มีรูปแบบตัวอักษรมาตรฐานตามคุณลักษณะตัวอักษรภาษาไทยแบบดั้งเดิมที่ประกอบไปด้วย เส้นตรง เส้นโค้ง เส้นซิกแซก และการเริ่มต้นตัวอักษรด้วยลูปวงกลมหรือที่เรียกว่าหัวตัวอักษร ส่วนกลุ่มรูปแบบตัวอักษรที่โครงสร้างของตัวอักษรปราศจากส่วนหัวตัวอักษรเช่น “ก” “ฅ” “ท” “ด” “ต” “บ” “ป” “ข” “ช” “ซ” จากตัวอย่างดังกล่าว ตัวอักษรปราศจากส่วนหัวตัวอักษรนั้นจะเห็นว่าตัวอักษรมีรูปแบบคล้ายกันมากขึ้น และปัญหาที่ตามมาคือความสามารถในการรู้จำตัวอักษรภาษาไทยรูปแบบตัวอักษรที่ไม่มีหัวยังไม่ถูกต้องเท่าที่ควร

ดังนั้นวิทยานิพนธ์นี้จึงได้นำเสนอแนวคิดวิธีการใหม่โดยใช้การแบ่งแยกตามลำดับชั้นในโครงข่ายประสาทเทียม Hierarchical Normalization Cross-Correlation ARTMAP สำหรับการรู้จำรูปแบบตัวอักษรพิมพ์ภาษาไทยที่ไม่มีหัว และนำเอาหลักการของครอสคอร์รีเลชันมาใช้เป็นฟังก์ชันสำหรับการวัดค่ามาตรฐานความคล้ายคลึงระหว่าง Input Vector กับ Weight Vector ข้อดีของการใช้ ครอสคอร์รีเลชันคือสามารถวัดค่าความคล้ายคลึงกันของรูปแบบในแต่ละส่วนออกมาได้โดยไม่ขึ้นกับขนาด เพราะเป็นการเปรียบเทียบความเหมือนในแต่ละส่วนของตัวอักษร โดยไม่จำเป็นต้องมีขนาดเริ่มต้นที่เท่ากัน กลุ่มข้อมูลที่ใช้ในการ Train และ Test ประกอบด้วยรูปแบบตัวอักษรพิมพ์ภาษาไทยไม่มีหัว 39 แบบและตัวอักษรบนภาพจากป้ายทะเบียนรถไทย 506 ป้าย ผลการทดลองที่ได้จากการรู้จำด้วยวิธีใหม่ที่น่าเสนอโดยเฉลี่ยเท่ากับ 88.49% นอกจากนี้ยังนำกลุ่มข้อมูลไปทดสอบกับโปรแกรม AmThai, ThaiOCR และ Fuzzy ARTMAP

The relevant researches concerned with Thai Character Recognition in the past have been emphasized in ability enhancement in the recognition rate by considering the structure of the standard character pattern based on the Feature Characteristic. Thai characters which are composed of direct lines, curves, zig-zag lines and the beginning of writing with the head circle loop which is known as the head of character. The examples of Thai Characters without the head circle loop are “ก” “ข” “ค” “ด” “ต” “อ” “ป” “ผ” “บ” “พ” “จ” “ฉ” “ช”. From these examples, it can be seen that each single character is much more similar to each other. The following problem is that the ability for recognition rate of non-head-loop is still not on the right way.

Therefore, this thesis presents the new methodology utilizing the level classification and Hierarchical Normalization Cross-Correlation ARTMAP for recognizing Thai Characters that have pattern without head circle loop. This research applied with the cross-correlation method which is used to be a criteria function to measure the similarity between the input vector and the weight vector. The advantage of using cross-correlation function is that the similarity of the pattern in each part can be independently measured concerning the size of character because the method is to compare the similarity of each character regardless the same size of starting. The set of data used in training and testing sets compose of non-circle head loop 39 fonts and the characters from 506 vehicle license plates. The result from this new method is shown the average of 88.49 percent recognition rate. Moreover, the data set has been investigated by testing with the program AmThai, ThaiOCR, and Fuzzy ARTMAP.